**Declaration**

I hereby declare that the project entitled MedScrape: To Irradicate Common Medication Mistakes submitted by us to the School of Computer Science and Engineering, VIT Chennai, 600127 in partial fulfilment of the requirements of the award of the degree of Bachelor of Technology in Computer Science and Engineering is a bona-fide record of the work carried out by me under the supervision of Dr. Gayathri R. I further declare that the work reported in this project, has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma of this institute or of any other institute or University.

Place: Chennai

Date: 29/04/2022

Signature of the candidate

**Certificate**

This is to certify that the report entitled MedScrape: To Irradicate Common Medication mistakes is prepared and submitted by us to VIT Chennai, in partial fulfilment of the requirement for the award of the degree of Bachelors of Technology in Computer Science and Technology is a bona-fide record carried out under my guidance. The project fulfils the requirements as per the regulations of VIT and in my opinion meets the necessary standards for submission. The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma and the same is certified.

**Acknowledgement**

I obliged to give my appreciation to a number of people without whom I could not have completed this project successfully.

I would like to place on record my deep sense of gratitude and thanks to my internal guide Prof. Gayathri R, School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology, Chennai, whose esteemed support and immense guidance encouraged me to complete the project successfully.

I would like to thank our HoD Dr. P. Nithyanandam, School of Computer Science and Engineering (SCOPE) and Project Co-Ordinator Prof. Gayathri R, Vellore Institute of Technology, Chennai, for their valuable support and encouragement to take up and complete this report.

I thank our management of Vellore Institute of Technology, Chennai, for permitting me to use the library and laboratory resources. I also thank all the faculty members for giving me the courage and the strength that I needed to complete my goal. This acknowledgment would be incomplete without expressing the whole hearted thanks to my family and friends who motivated me during the course of my work.

**Abstract**

Medscrape is used to irradicate common medication mistakes people do and not adhering to the medication, People usually consume medicines that is not to be consumed for the symptoms they are facing which may lead to serious health problems and even cost his/her life.

It also tells the patient/user when and in case of what symptoms they should consume the medicine they want to know about. It can also be used by people who tend to forget about the medicine they have to consume and thus can use our website to upload the image of the medicine that they don’t know about it and get the correct information regarding it. Our project helps in classification of the medicine image that user wants to know about and return him/her the name of the medicine detected and provide the required information about the medicine which is medicine name, introduction, side effect and when to consume. The image classification is done using Transfer Learning model InceptionV3 which processes the medicine image given by the patient/user on our website and provides the user with the correct information which is rendered on the web page.

**Introduction**

This application helps to classify different medication pills. There have been numerous cases of people falling a lot sicker than before and numerous cases of deaths because of wrong medication. Because of this even doctors have been wrongly accused of deaths as well. It has been a major problem but still has been unaddressed. With this we are trying to reduce the wrong taking of pills.

This project consists of 3 major sections:

Front-End:

Nodejs has been used which an open-source, cross-platform, back-end JavaScript runtime environment that runs on the V8 engine and executes JavaScript code outside a web browser. Node.js lets developers use JavaScript to write command line tools and for server-side scripting—running scripts server-side to produce dynamic web page content before the page is sent to the user's web browser.

Web Scraping:

Even for web scraping we used Nodejs

Back-End:

We used Django which is a high-level Python web framework that enables rapid development of secure and maintainable websites. Built by experienced developers, Django takes care of much of the hassle of web development, so you can focus on writing your app without needing to reinvent the wheel.

Transfer Learning:

For transfer learning we used the pretrained model InceptionV3

Virtualisation:

Microsoft Azure has been used for hosting the website which is a **Microsoft Azure**, often referred to as **Azure** is a [cloud computing](https://en.wikipedia.org/wiki/Cloud_computing) service operated by [Microsoft](https://en.wikipedia.org/wiki/Microsoft) for application management via Microsoft-managed [data centers](https://en.wikipedia.org/wiki/Data_center). It provides [software as a service (SaaS)](https://en.wikipedia.org/wiki/Software_as_a_service), [platform as a service (PaaS)](https://en.wikipedia.org/wiki/Platform_as_a_service) and [infrastructure as a service (IaaS)](https://en.wikipedia.org/wiki/Infrastructure_as_a_service) and supports many different [programming languages](https://en.wikipedia.org/wiki/Programming_language), tools, and frameworks, including both Microsoft-specific and third-party software and systems.

**Planning and Requirement Specifications**

**Literature Review:**

In this section the discussion on similar research paper and publication published.

* [1] Proposed an automatic classification system for pill images based on their shape and color. using image processing techniques to specify an attribute set used by Support Vector Machines and Multilayer Perceptron classifiers.
* [2] Providing the right treatment plan for patients includes knowledge about their current medications and drug allergies, an often-challenging task. The widespread growth of prescribing and consuming medications has increased the need for applications that support medication reconciliation. They show a deep-learning application that can help reduce avoidable errors with their attendant risk, i.e., correctly identifying prescription medication, which is currently a tedious and error-prone task.
* [3] The physical form of the medication, often tablets and capsules, captures the unique features of the NDC (NATIONAL DRUG CODE) product to help ensure patients receive the same medication product inside their prescription bottle as is found on the label from a pharmacy. An automated check is used to report, evaluate and predict the shape, color, and NDC for images showing a pile of pills inside a prescription bottle. Patterns of incorrect NDC predictions based on similar colors, shapes, and imprints of pills were identified and recommendations to improve the model were provided.

**Dataset**

For now, we have taken 4 medicinal pills:

* Allegra
* Antacid
* Paracetamol
* Statin

For training and we scraped the data from Google Images. 50 images per medicine has been scraped for each medicine and 10 images each for testing.

Requirements

3.1 System Requirements

3.2 Hardware Requirements

* Operating system windows (7 0r later version)
* Ram 8 GB
* Hard disk Space- 100 GB
* Processor- Intel i3 1.8 gHz or newer versions

2.3.3 Software Requirements

Language:

* Python
* Node Js
* Django
* Putty

Enviroment:

* Anaconda Distribution
* VS Code

Packages Used:

* ﻿﻿absl-py==0.12.0
* asgiref==3.3.4
* astunparse==1.6.3
* cachetools==4.2.2
* certifi==2020.12.5
* chardet==4.0.0
* Django==3.2.3
* django-cors-headers==3.7.0
* djangorestframework==3.12.4
* flatbuffers==1.12
* gast==0.4.0
* google-auth==1.30.0
* google-auth-oauthlib==0.4.4
* google-pasta==0.2.0
* grpcio==1.34.1
* h5py==3.1.0
* idna==2.10
* keras-nightly==2.5.0.dev2021032900
* Keras-Preprocessing==1.1.2
* Markdown==3.3.4
* numpy==1.19.5
* oauthlib==3.1.0
* opt-einsum==3.3.0
* Pillow==8.2.0
* protobuf==3.17.0
* pyasn1==0.4.8
* pyasn1-modules==0.2.8
* pytz==2021.1
* requests==2.25.1
* requests-oauthlib==1.3.0
* rsa==4.7.2
* six==1.15.0
* sqlparse==0.4.1
* tensorboard==2.5.0
* tensorboard-data-server==0.6.1
* tensorboard-plugin-wit==1.8.0
* tensorflow==2.5.0
* tensorflow-estimator==2.5.0
* termcolor==1.1.0
* typing-extensions==3.7.4.3
* urllib3==1.26.4
* Werkzeug==2.0.0
* wrapt==1.12.1

**Module Explanation**

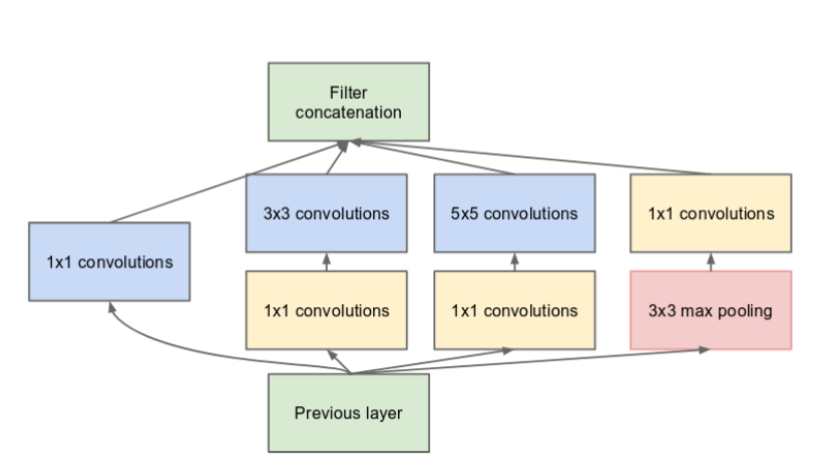
1. Front-End

A very basic but unique website has been created for the ease use. This website has two web pages. One for uploading the medicine and as processing the result and another for the users for uploading the details of the medicine for which they want us to display but is not our dataset yet.

2. Backend

This part helps to connect the front end and the transfer learning. Whenever the user inputs the picture, it is then provided to the backend and is then taken as input to the model. The output received by the model is then shown on the website as per the search query.

3. Transfer Learning model: Inception V3



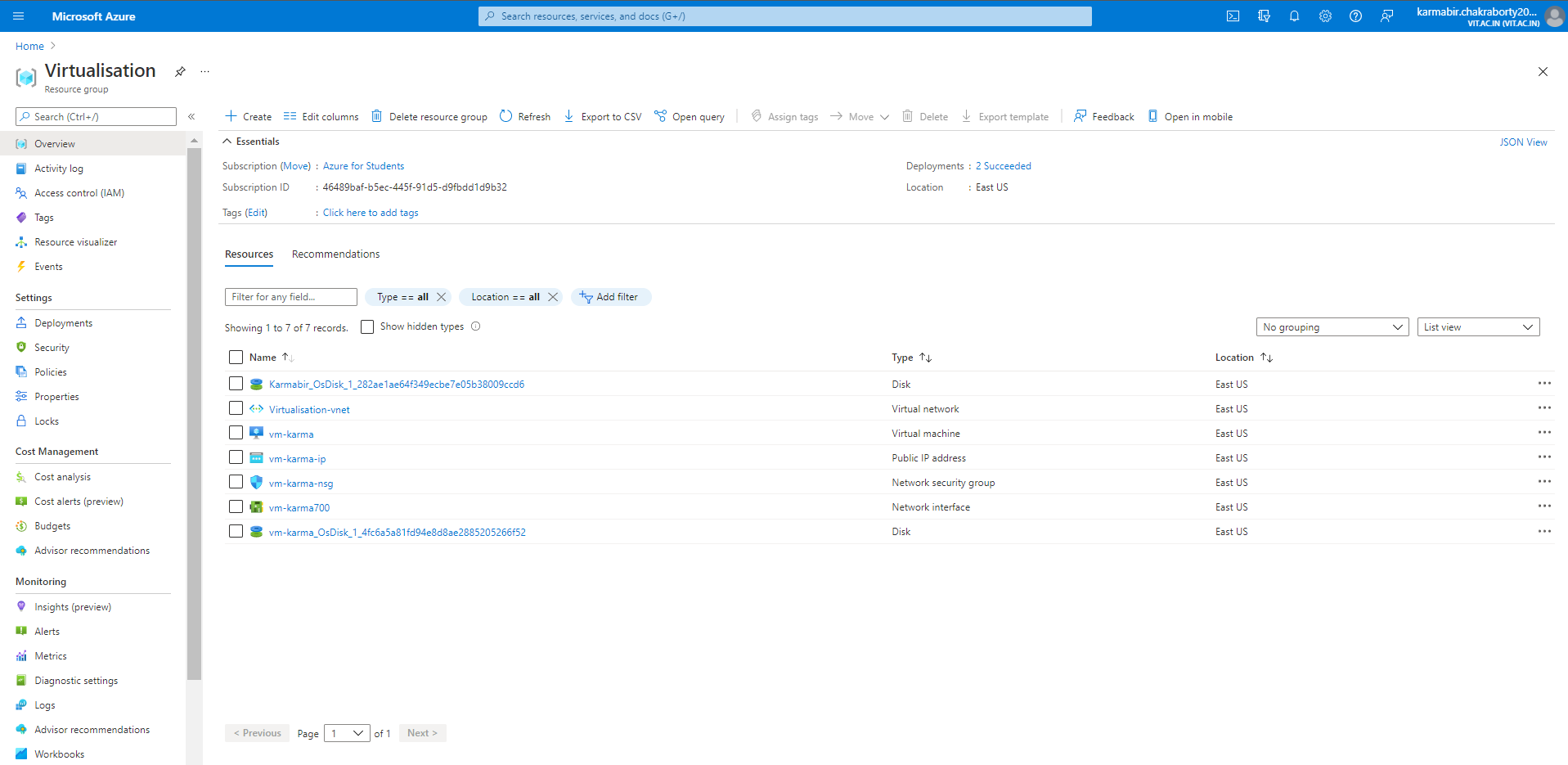
Inception V3 by Google is the 3rd version in a series of Deep Learning Convolutional Architectures. Inception V3 was trained using a dataset of 1,000 classes (See the list of classes [here](https://gist.github.com/yrevar/942d3a0ac09ec9e5eb3a)) from the original ImageNet dataset which was trained with over 1 million training images, the Tensorflow version has 1,001 classes which is due to an [additional "background' class](https://github.com/tensorflow/models/blob/master/research/slim/datasets/imagenet.py#L83) not used in the original ImageNet. Inception V3 was trained for the ImageNet Large Visual Recognition Challenge where it was a first runner up.

4. Virtualisation:

We created a virtual machine on Microsoft Azure where we installed all the libraries, linked with our github repository where put up all our code. We finally held our backend there and were able to run it. The steps for implementation for virtualising or hosting the application are provided in the next section.

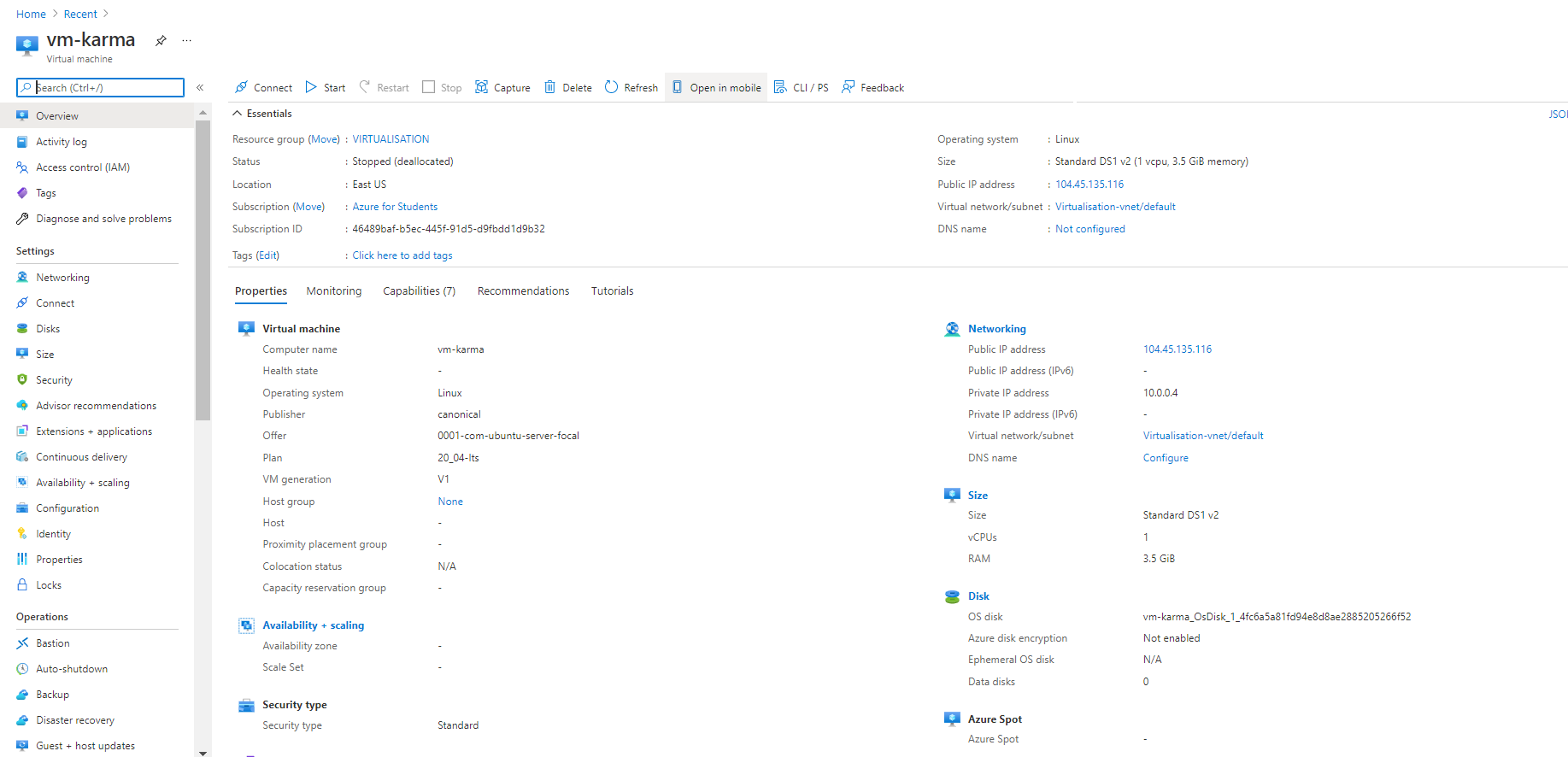
**Implementation**

* Whole resource group:

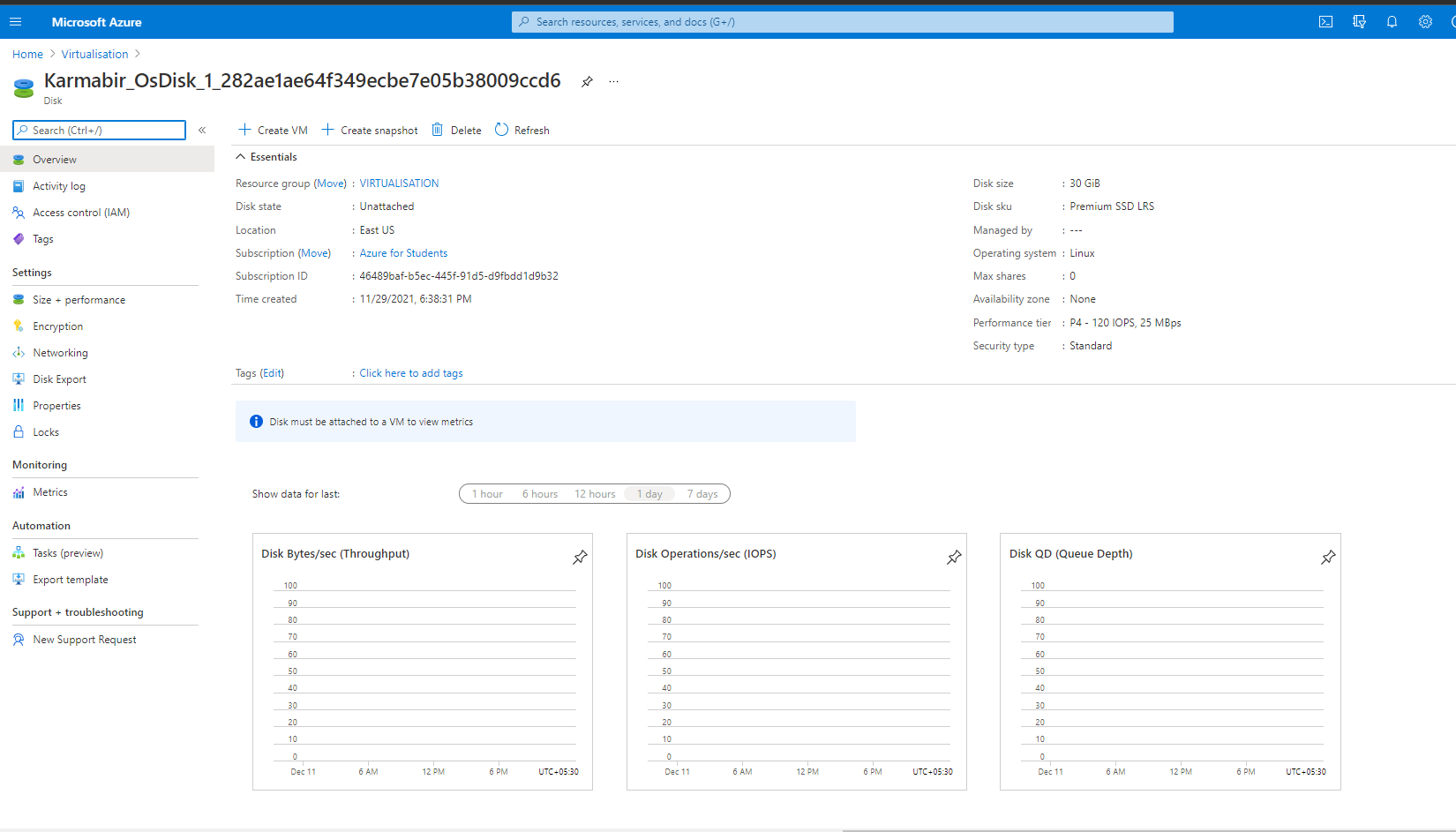


Resources Created:

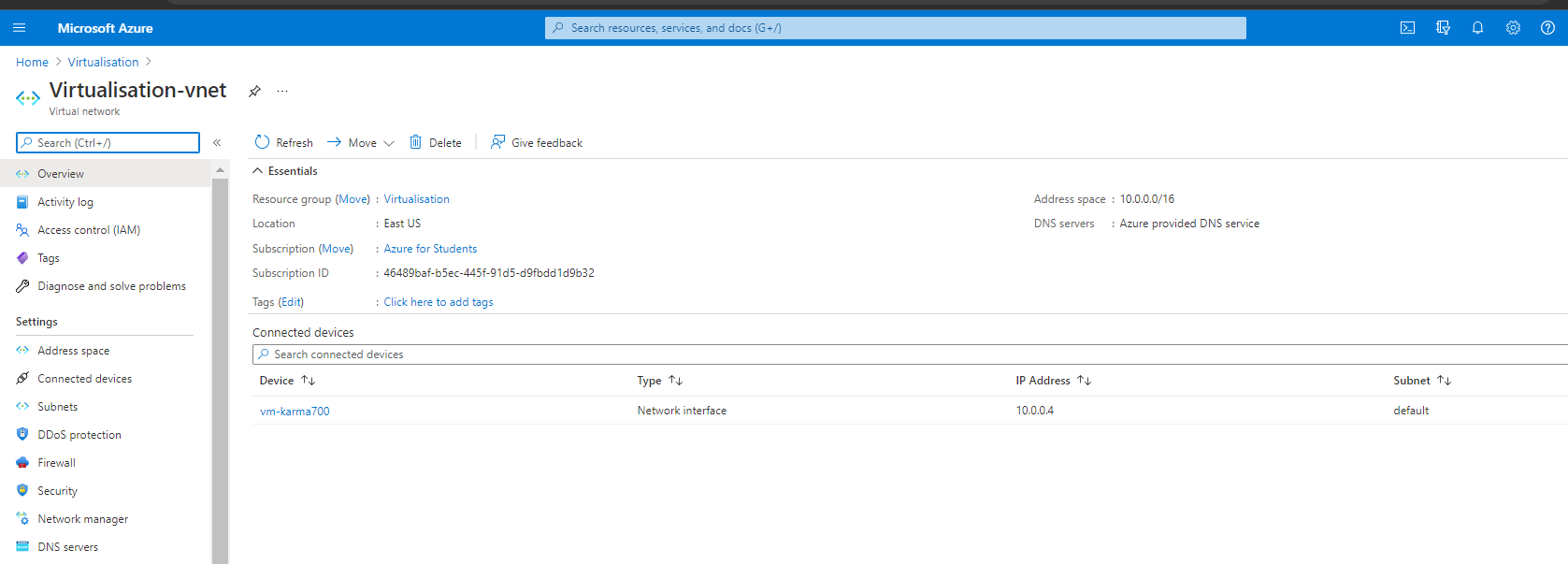
* Virtual machine:



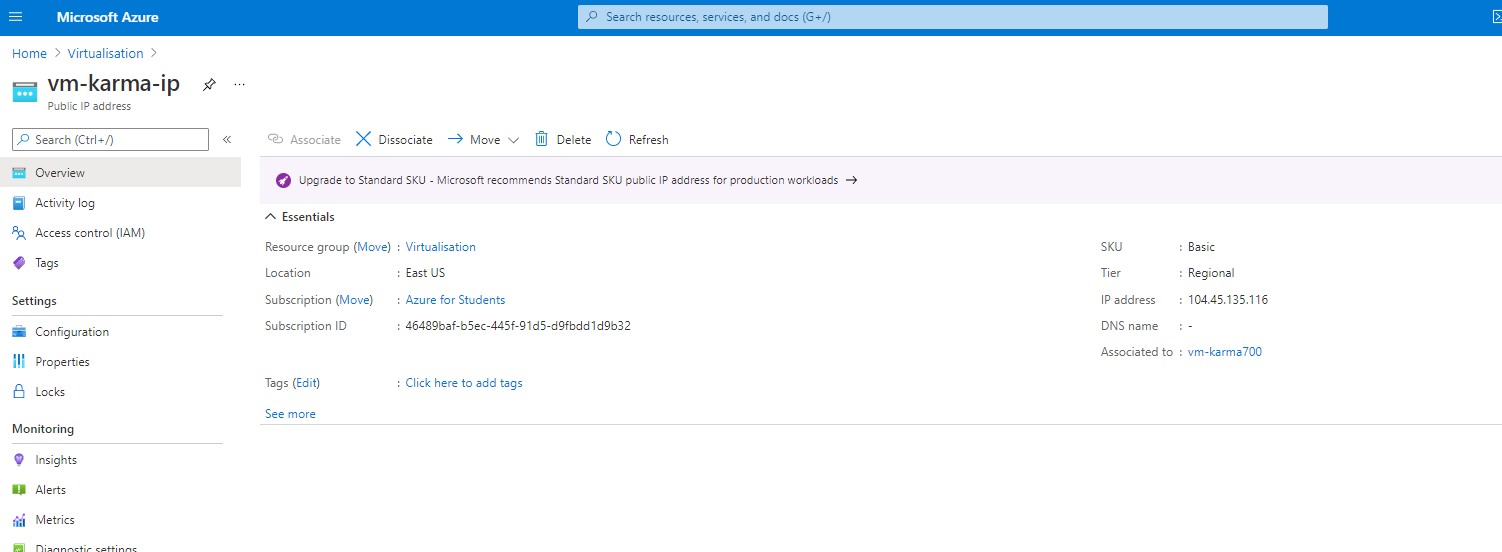
* OS Disk:



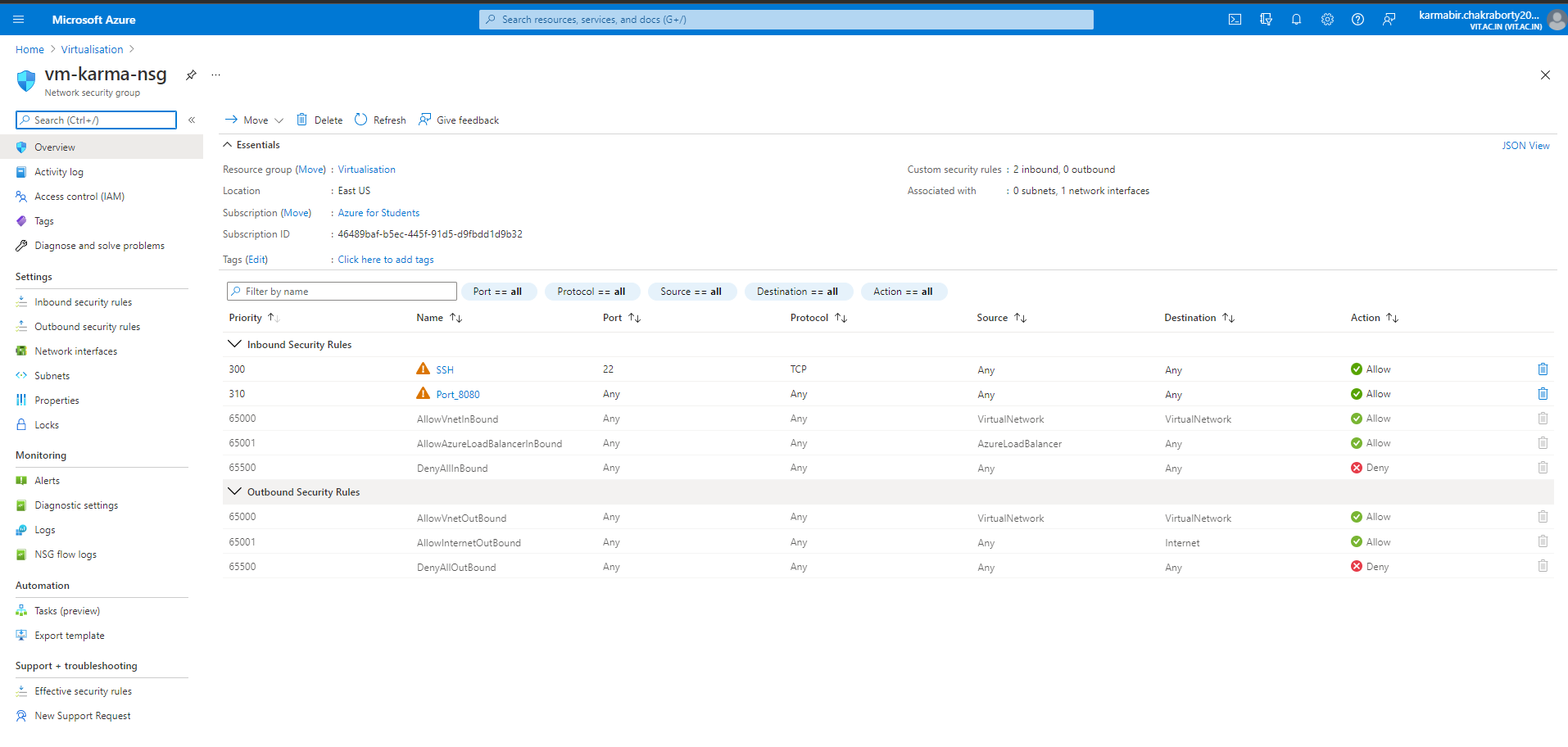
* Virtual Network:



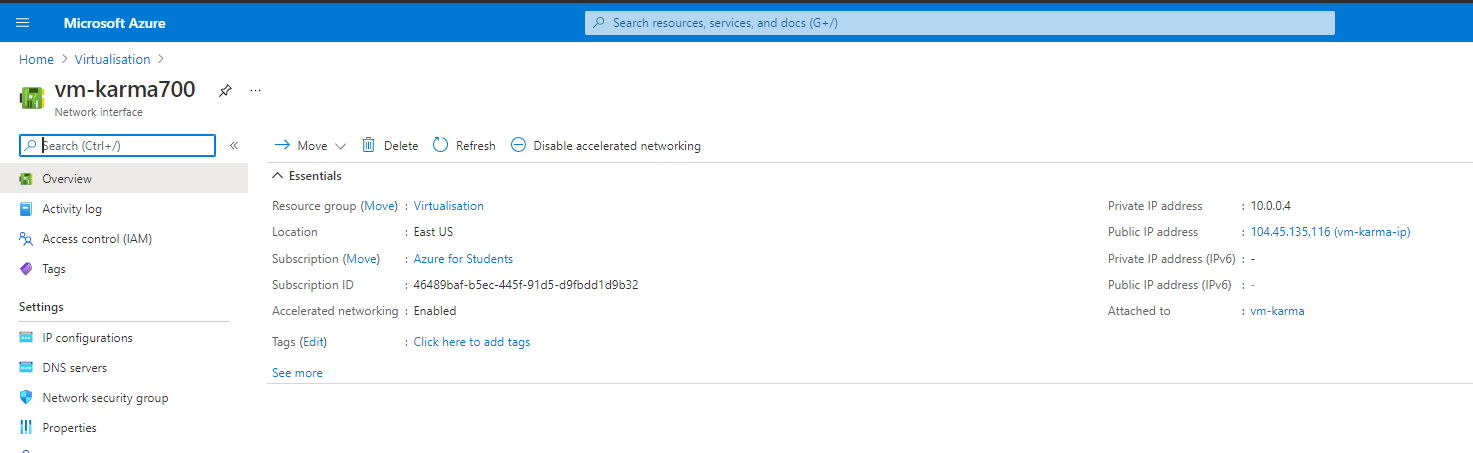
* Public IP address:



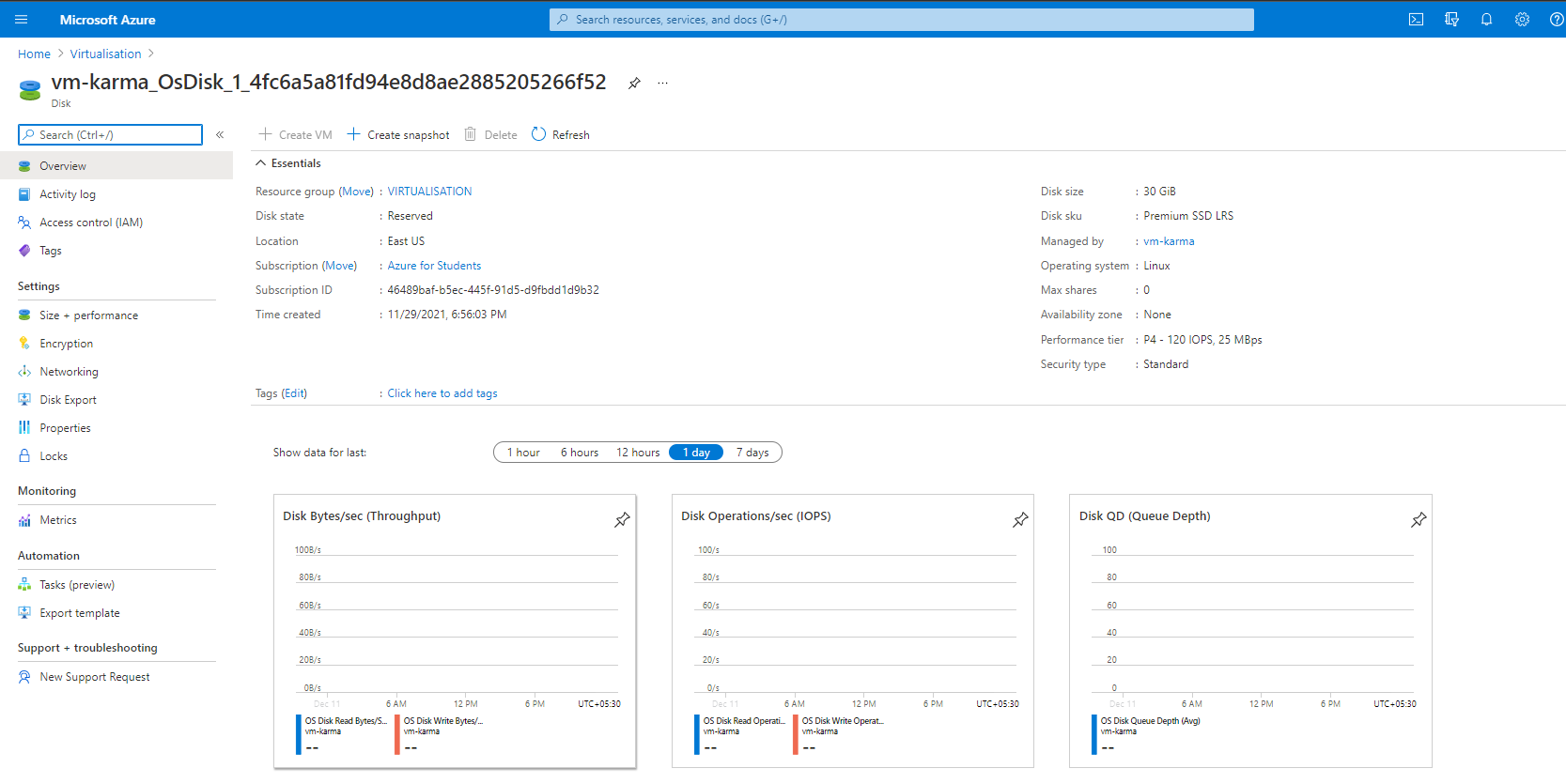
* Network Security Group:



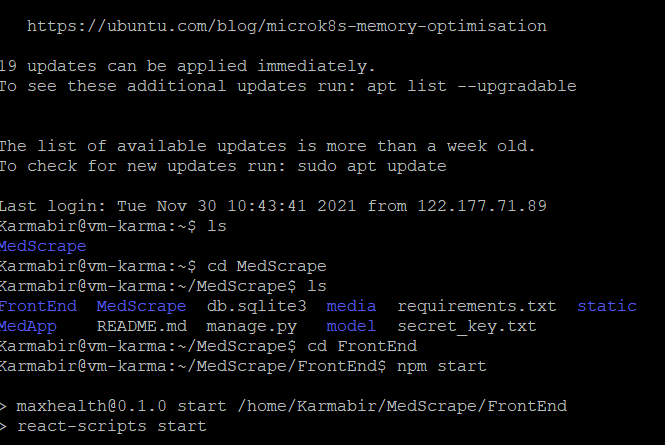
* Network Interface:



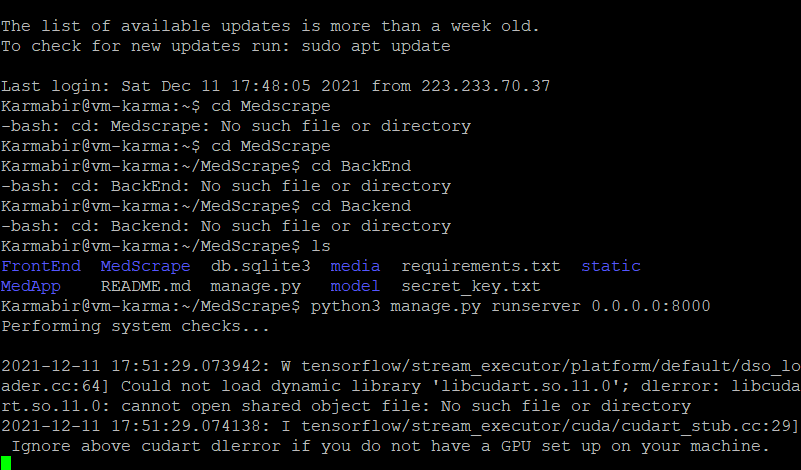
* OS Disk-2:



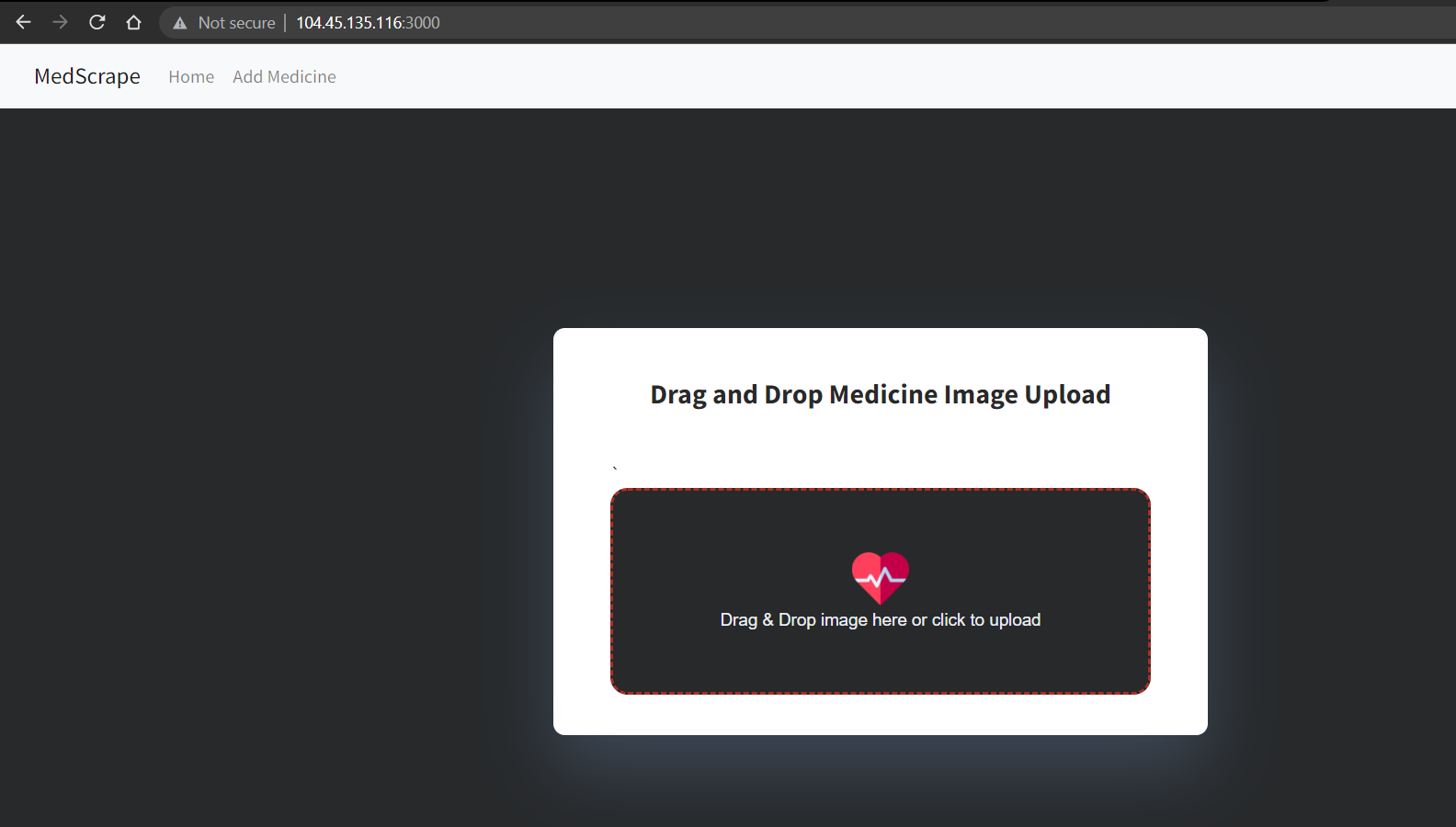
* Putty:
* For frontend:

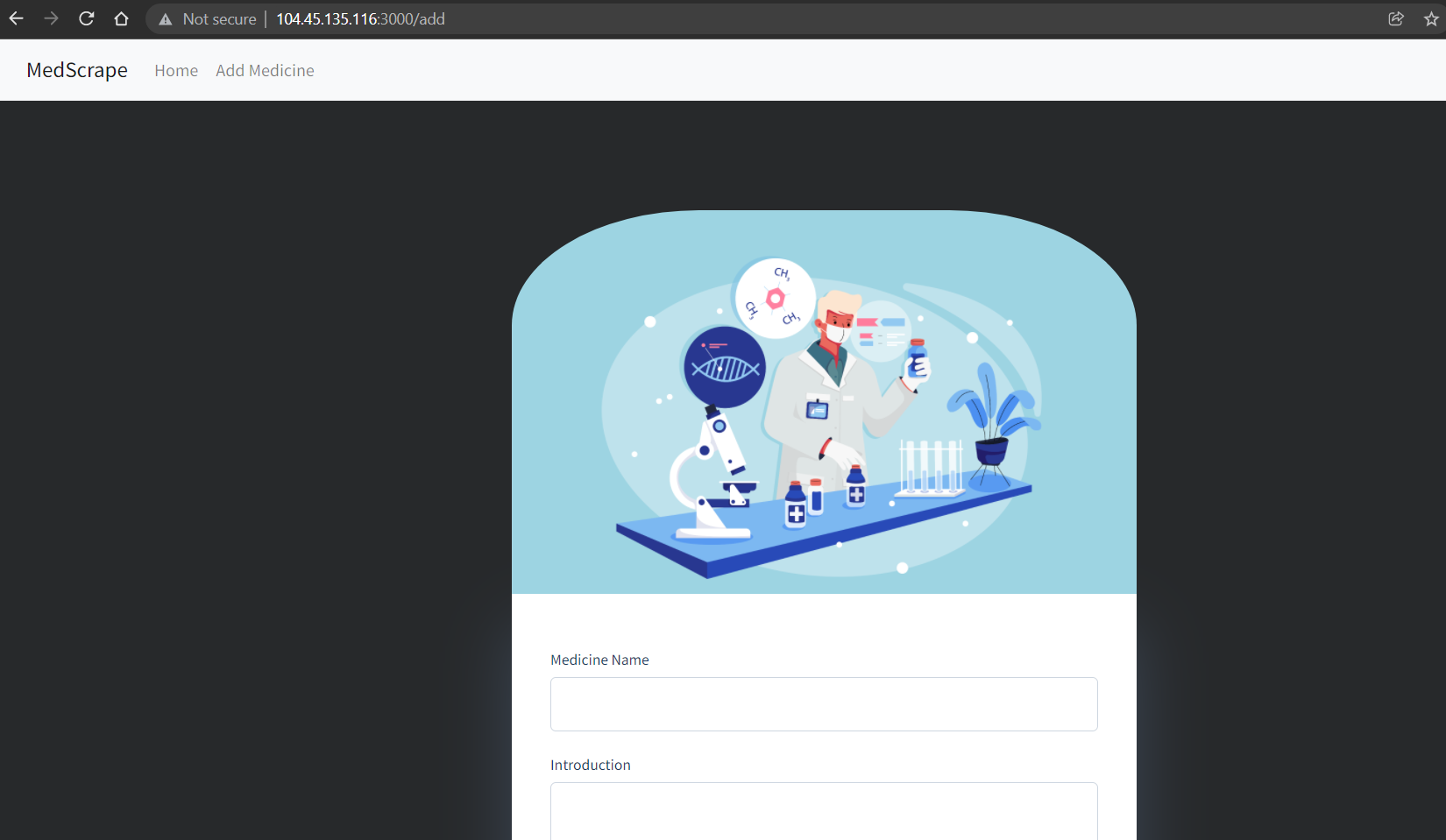


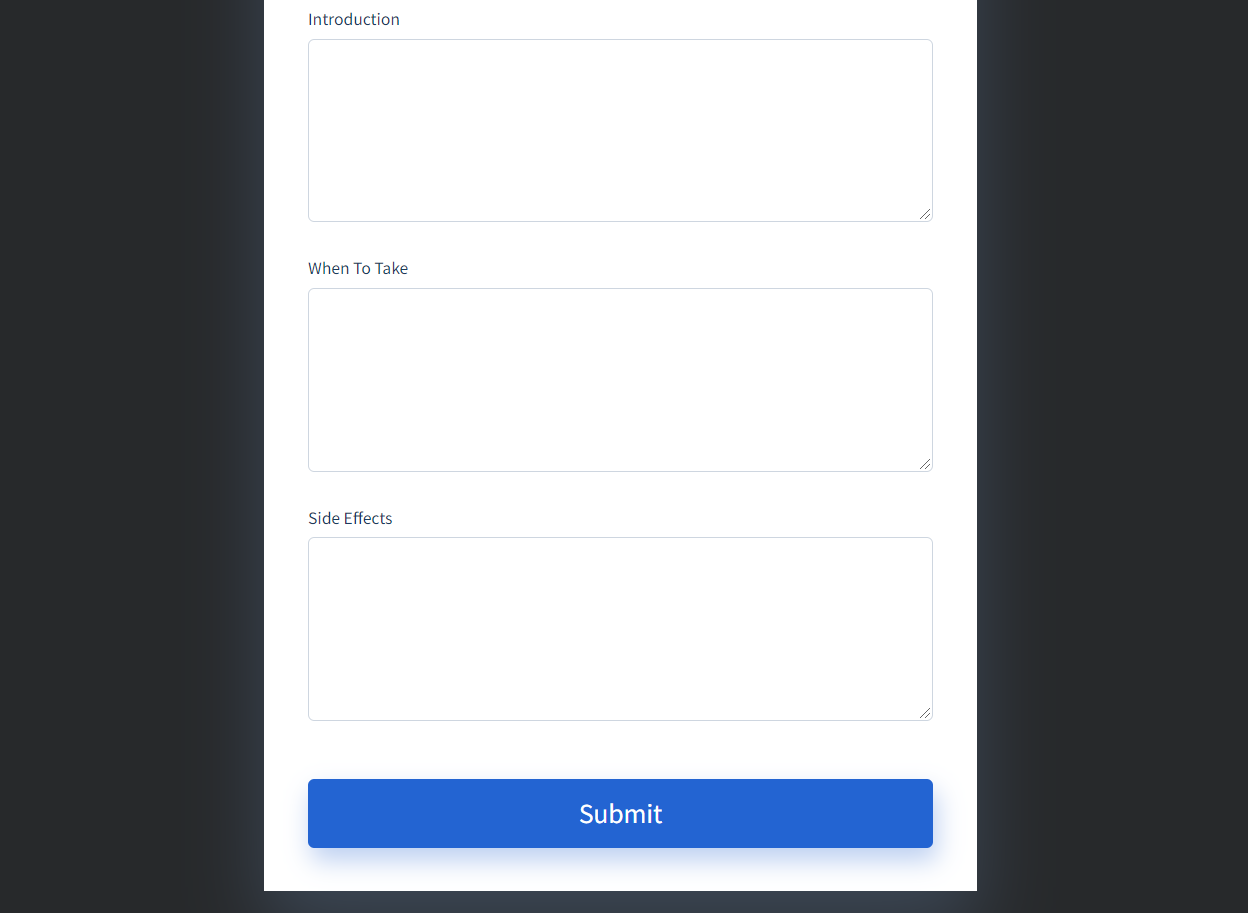
* Backend:



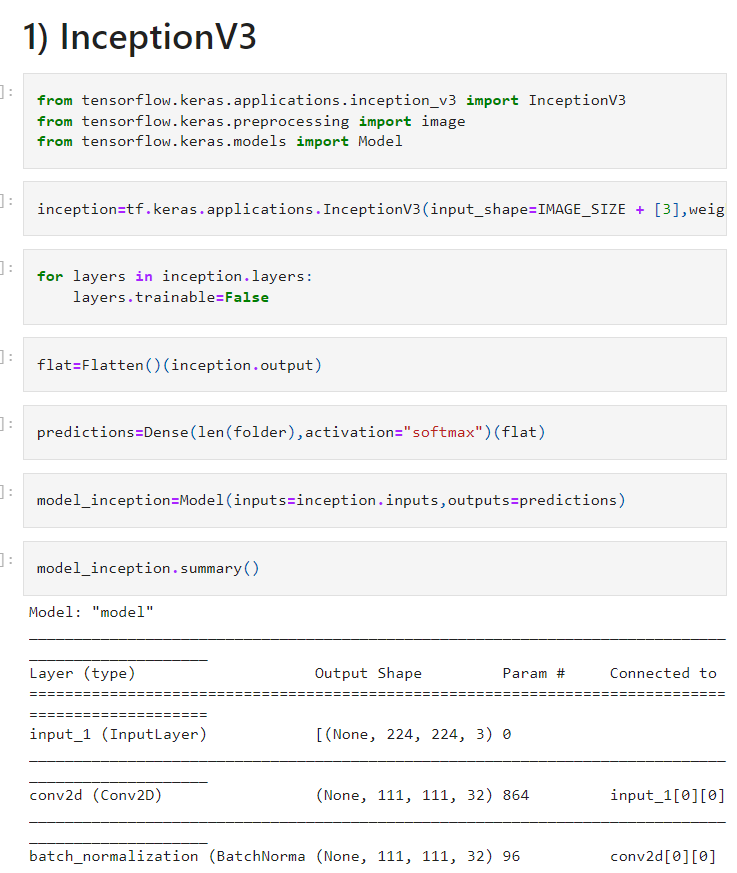
* Website Frontend:

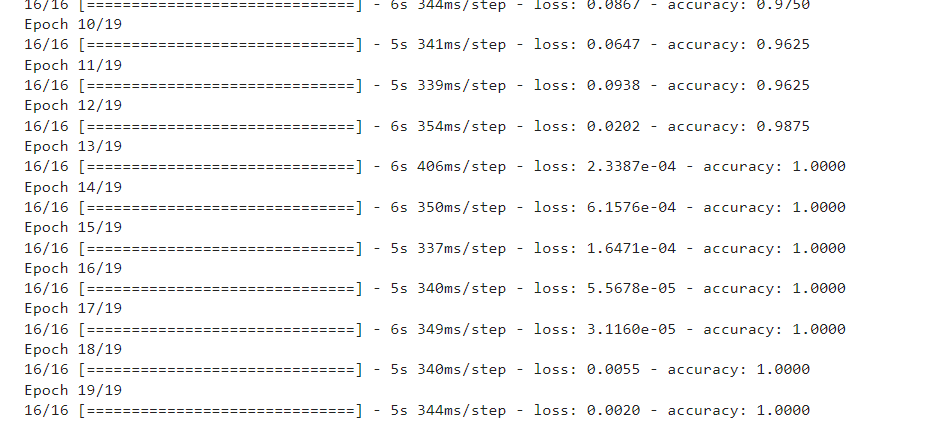






* Model: 









**References**

[1] Pill Image Classification using Machine Learning2019 8th Brazilian Conference.

[2] Natalia Larios Delgado, Naoto Usuyama, Amanda K. Hall, Rebecca J. Hazen, Max Ma, Siva Sahu & Jessica Lundin npj Digital Medicine volume 2, Article number: 10 (2019).

[3] Performance evaluation of a prescription medication image classification model: an observational cohortCorey A. Lester, Jiazhao Li, Yuting Ding, Brigid Rowell, Jessie ‘Xi’ Yang & Raed Al Kontar npj Digital Medicine volume 4, Article number: 118 (2021).

L. T. Kohn, J. Corrigan, M. S. Donaldson et al., To err is human: building a safer health system. National Academy Press Washington, DC, 2000, vol. 6. J. A. Greene and A. S. Kesselheim, “Why do the same drugs look different? pills, trade dress, and public health,” New England Journal of Medicine, vol. 365, no. 1, pp. 83–89, 2011, pMID: 21732842.